

Assessment of usability of artificial intelligence (AI) and recommended procedure for evaluations and impact assessments (IA)

Artificial intelligence (AI) is increasingly penetrating many human activities, including evaluation and impact assessments.

Evaluators use a whole range of tools in order to evaluate the feasibility, the status, or the progress of a project, a program, a strategy, or even the implementation of a policy. This involves processing "tons of" information that comes in all sorts of different formats: documents of different kinds, scientific literature, interview transcripts based on different interviewing techniques, translated texts, intercultural communications, workshop discussions and reports, surveys of different quality, monitoring databases, analysis of systems, and so on. So how tempting would it not be to just have our ICT devices draw on artificial intelligence to chew through and digest all of this material, organize it for us, double check it against all of the clutter available in the internet of thoughts, and churn out a draft report that we only need to corroborate in the field, or merely sign off and collect our fees.

Or does this just seem too good to be true?

Protean Consulting's Summary View of the Use of Artificial Intelligence in Evaluations

Like any other potential tool, AI must be assessed based on its capabilities and merits as well as on its disabilities and shortcomings. It does not enjoy any privileged status in comparison with other tools, many of which have proven their merit and have a time tested record of successful applications.

Important things to know about artificial intelligence

- A) AI is a fast evolving technology with both upsides and downsides, the reliability of which is unclear. Evaluators and impact assessors **should adhere to scientific standards which currently approve AI usage only for refinement of language used**. Evaluators and impact assessors may resort to current text editing tools as an alternative to AI.
- B) Only private AI can provide a meaningful assurance of confidentiality. **Any use of open AI tools via the internet implies a breach of confidentiality that cannot be prevented**.

- C) Private AI **requires significant upfront investment** (hardware such as GPU, not only the users' devices), including extensive (i.e. energy and time intensive) training of the respective models **before they can be used for the specific evaluation or impact assessment**.
- D) Should the client of the evaluation or impact assessment offer the use of its own private AI systems, caution is warranted because the AI systems of the client are bound to be trained in line with client preferences and needs. Consideration of any **AI generated content by client systems is to be taken with the appropriate grain of salt. Any confidential data** collected during an evaluation or impact assessment **must not be transferred to the clients' systems**.
- E) AI data is not the information evaluators and impact assessors evaluate and assess. It is **project data and information** that is being evaluated and assessed.

Important things to do when dealing with artificial intelligence in evaluations and impact assessments

- A) Give time to the involved stakeholders to individually think about the **usability and implications of the use of AI** in the evaluation or impact assessment.
- B) Organize a process by which the merits of any potential AI tools can be jointly **assessed, discussed, and agreed upon by all** involved stakeholders (e.g. workshop)
- C) Document the results of the stakeholder workshop and ensure that the documentation is approved by all of the stakeholders (i.e. **formally required confirmations**) and distribute the final document to all stakeholders.
- D) Be **prepared to review and revise** the agreement reached during implementation if any concerns relating to the ethical use, the quality of results, or any other important (possibly unintended) results of the tools in usage are being raised.
- E) Adjust your evaluation or impact assessment implementation plan in line with the outcome of step 4 **even if it delays the process** because credibility of the evaluation or impact assessment overrides all other concerns.

Dr Sylvia CHUNG Mei Ying
Christoph David WEINMANN

Kuching, Sarawak, Malaysia, 15/08/2025
[my-protean.xyz]